Running head: IIV MODELING

Novel Full-Bayesian and Hybrid-Bayesian Approaches for Modeling Intraindividual Variability

Yuan Fang & Lijuan Wang

Department of Psychology

University of Notre Dame

Lijuan Wang is grateful for the support from IES grant R305D210023 during the study.

Correspondence should be addressed to Lijuan Wang, Department of Psychology, University of Notre Dame, Notre Dame, IN, 46530. Email: lwang4@nd.edu.

Abstract

Intraindividual variability (IIV) characterizes the amplitude and temporal dependency of short-term fluctuations of a variable and is often used to predict outcomes in psychological studies. However, how to properly model IIV is understudied. In particular, intraindividual standard deviation (or variance), which quantifies the amplitude of fluctuation of a variable around its mean level, can be challenging to model directly in popular latent variable frameworks, such as dynamic structural equation modeling (DSEM). In this study, we introduced three novel modeling methods, including two two-step hybrid-Bayesian methods using DSEM and a one-step full Bayesian method, to model IIV as predictors. We conducted a simulation study to evaluate the performance of the three methods and compared their performance to that of the conventional regression approach under various data conditions. Simulation results showed that the hybrid-Bayesian approach with multiple draws (HBM) and the one-step full Bayesian (FB) approach performed well in recovering the parameters when sufficient sample size and time points were available. The data requirement of using FB was lower than HBM. However, the conventional approach and hybrid-Bayesian approach with a single draw failed to recover parameters, even with large samples. We provided a simulated data example with code online to illustrate the use of the methods. *Keywords*: intraindividual variability, dynamic structural equation modeling, Bayesian estimation, multiple imputation

Novel Full-Bayesian and Hybrid-Bayesian Approaches for Modeling Intraindividual Variability

Introduction

Nesselroade (1991) distinguished the concepts between intraindividual variability (IIV), which characterizes short-term fluctuations of a variable around its mean level that are "more or less reversible and occur rapidly" (p. 215), and intraindividual change, which is enduring, developmental, and irreversible. Although traditionally treated as random noise, IIV was found to have substantive meanings, such that individuals may consistently exhibit similar or different patterns of intraindividual variability (Nesselroade, 1991, 2001; Nesselroade & Molenaar, 2010). These pioneering work motivated the study of interindividual differences in intraindividual variability in many areas of psychology. For example, in aging studies, intraindividual variability in reaction time tasks of older adults was found to be negatively related to their cognitive performance over three years (Bielak et al., 2010). In affect dynamics research, Ong et al. (2025) found that higher mean levels and inertia of negative effect were associated with depressive symptoms. Furthermore, in the personality domain, Wright & Jackson (2025) found that intraindividual variability around personality trajectories predicted health status using big longitudinal data from five countries.

In the last several decades, intensive longitudinal designs, such as the measurement burst design (e.g., Nesselroade, 1991; Ram & Gerstorf, 2009), experience sampling (e.g., Carstensen et al., 2011), and daily diary design (e.g., Laurenceau & Bolger, 2005), have gained popularity for collecting many repeated measurements over short periods of time, making it feasible to study interindividual differences in intraindividual variability. The focus of research has transitioned from whether to study intraindividual variability to how

to study it (Boker et al., 2009). Ram & Gerstorf (2009) categorized intraindividual variability into net (or time-unstructured) versus time-structured intraindividual variability. The key distinction is that net intraindividual variability reflects dynamic characteristics or time-unrelated change (e.g., the amplitude of fluctuations), whereas time-structured intraindividual variability describes dynamic processes or time-dependent change (e.g., cycles, oscillations, and lagged effects). More specifically, net intraindividual variability refers to the range or dispersion of behaviors exhibited by an individual, with little consideration of the temporal order of measurements. In contrast, time-structured intraindividual variability emphasizes temporal dependencies across occasions. Studying time-structured intraindividual variability allows researchers to better understand how individuals adapt to endogenous and exogenous influences over time.

To measure net intraindividual variability, a frequently used indicator is the intraindividual standard deviation (ISD), reflecting the extent to which an individual's scores vary around their own intraindividual mean score (Dejonckheere et al., 2019; Ram & Gerstorf, 2009). Alternative indicators include the intraindividual variance (IVAR), absolute range, interquartile range, and others. Among these, IVAR is often used, because its distribution is easier to model than the others (Wang et al., 2012). Time-structured intraindividual variability is typically assessed using time series models such as the autoregressive models or moving average models, capturing systematic, time-dependent structures in the data. To quantify time-structured intraindividual variability, the first-order autocorrelation and the first-order autoregressive parameter (i.e., the AR(1) coefficient, also called the inertia) are commonly used in psychological research (Du & Wang, 2018). For example, the inertia reflects the extent to which a state at one time point can be predicted by its value at the previous time point.

Some IIV indicators, such as the mean square successive difference (MSSD; Jahng et al., 2008), combine both the amplitude of fluctuation and temporal dependency. However,

such composite measures may obscure the distinct contributions of each dynamic component when predicting an outcome. For example, Wang et al. (2012) argued that the amplitude of fluctuation and temporal dependency can have different predictability of outcomes and their real data analysis results confirmed this. Therefore, they recommended that the amplitude of fluctuation and temporal dependency represent distinct aspects of intraindividual variability, and thus should be modeled by separate indicators.

As reviewed earlier, there is much research interest in examining whether IIV (both the amplitude of fluctuations and the temporal dependency) is predictive of psychological health outcomes (e.g., Bielak et al., 2010; Ong et al., 2025; Wright & Jackson, 2025), allowing researchers to better understand interindividual differences and prediction of well-being and health (e.g., Kuppens et al., 2012). To better understand the current statistical practices in IIV research, we conducted a literature review on how intraindividual variability was quantified and modeled in psychological studies. We used PsycArticles with the keyword "variability" in the abstract to identify IIV articles published between 2010 and 2024 and in two journals: *Emotion* and *Psychology and Aging*. 81 studies were identified (see Appendix A in the Supplementary Materials for details). We extracted the following information from these studies, including the quantification indicators of IIV, whether IIV was modeled as a predictor in the analysis, and the modeling method of IIV (i.e., via latent variable or using computed indicators based on observed scores).

Among the 81 studies, 50 (61.7%) used ISD (or IVAR) to quantify intraindividual variability and 12 (14.8%) used AR(1) coefficients to study inertia. The rest of the articles used other indicators, such as MSSD (34.6%) or intraindividual coefficient of variation (ICV; 4.9%), to quantify intraindividual variability. In 43 (53.1%) of the 81 studies, intraindividual variability was included as a predictor in the analyses. In these studies, researchers are often interested in studying whether and how IIV indicators (e.g., ISD,

IVAR, MSSD, and/or AR coefficients) predict health and psychological outcomes, above and beyond the intraindividual mean. From the review, we also found that the prevalent method of modeling IIV was using IIV estimates directly calculated from the observed data. We refer to these IIV estimates as the observed IIV such as the observed ISD, observed IVAR, and observed AR(1) coefficient. For example, among the 50 studies that studied ISD (or IVAR), 44 (88%) modeled the observed ISD (or observed IVAR) as a predictor or outcome in the analyses. For inertia, out of the 12 studies, 10 (83.3%) used the observed AR(1) coefficient and modeled them in the next-step analyses. Due to its widespread use in current practice, we refer to the modeling of observed IIV as the conventional approach (or the "naive" approach).

However, researchers have noted a concern pertaining to the reliabilities of those IIV estimates, which can be computed using the squared correlation between the IIV indicator estimates and true values (e.g., between the observed ISD and the true ISD). For example, Estabrook et al. (2012) conducted a simulation study and found that the observed ISD had much lower reliability estimates compared to the observed intraindividual mean (IM) when the number of measurement occasions is small. Wang & Grimm (2012) analytically derived the reliabilities of the observed IM, observed ISD, and observed IVAR, and reported consistent findings. They also found that the observed IVAR can have low reliability estimates when the number of measurement occasions is small. Du & Wang (2018) extended this work and estimated the reliabilities of the observed IM, observed ISD, observed IVAR, and observed AR(1) coefficient with autocorrelated data. Notably, they found that the reliability estimates of the observed AR(1) coefficient were even lower than those of the observed ISD and observed IVAR under the same data conditions. In the conventional approach, these observed intraindividual variability indicators are directly modeled as predictors, which can lead to incorrect statistical inference about the regression coefficients, for instance, underestimating their relationships with outcomes (McDonald,

2013). To account for the reliability issue, researchers have proposed using the latent variable modeling or mixed-effects modeling frameworks in which intraindividual variability is modeled as latent variables or random effects. Below, we first review latent variable or mixed-effects approaches for modeling ISD and IVAR, and then turn to the approaches for modeling the inertia.

Gerstorf et al. (2009) first used item parceling to obtain three ISD estimates. Then, in a structural equation model (SEM), a latent ISD factor was constructed based on the three ISD estimates via the measurement model, and used to predict an outcome in the structure model. However, the performance of this approach has not been evaluated via analytical or simulation studies. Hedeker et al. (2008) proposed the mixed-effects location scale model (MELSM), where the IM (the location) and IVAR (the scale) are modeled as random effects, while allowing covariates to predict variations in both (Dzubur et al., 2020). Feng & Hancock (2024) extended MELSM to the structural equation modeling framework, enabling IM and IVAR to be modeled as latent predictors and/or latent outcomes. In these models, temporal dependency was not modeled, and thus time-structured intraindividual variability (e.g., the inertia) was not considered. In this case, the inertia is not included as a predictor of outcomes, although our literature survey suggested that researchers are interested in studying the predictivity of temporal dependency.

Rast et al. (2022) proposed a multivariate generalized autoregressive conditional heteroskedasticity model (MGARCH) to capture intraindividual variability through the shock variance (also called the residual or innovation variance), and to study the AR process in the shock variance for each individual. Dynamic structural equation modeling (DSEM; Asparouhov et al., 2018; Hamaker et al., 2018, 2021) was proposed to model the person-specific shock variance and AR parameters as latent predictors or outcomes at the between-person level. However, in these two models, it is challenging to directly model ISD

or IVAR as latent variables to study the amplitude of fluctuation in a process variable itself. Instead, the shock variance is used as an IIV indicator to capture dynamic characteristics. Importantly, when temporal dependency is present and modeled, the shock variance differs from the intraindividual variance (IVAR) of the process variable. Unlike ISD or IVAR, the shock variance depends on which dynamic process is modeled (Brockwell & Davis, 2002). For example, an AR(1) model and an AR(2) model can yield different shock variances using the same data. Additionally, as reviewed in our literature survey, researchers are more interested in examining individual differences in the intraindividual variability of the process itself rather than the residuals, and thus used ISD or IVAR as indicators of net intraindividual variability more frequently in the IIV literature (e.g., Dejonckheere et al., 2019; Ong et al., 2025; Wright & Jackson, 2025). Therefore, in the current study, we focus on using IVAR and ISD as indicators of net intraindividual variability indicators and study how to model them.

In the recently developed continuous time - residual DSEM (CT-RDSEM) framework (Asparouhov & Muthén, 2024a), IVAR rather than the shock variance can be directly modeled as a latent variable. However, the AR coefficients (e.g., the inertia) are not directly modeled in this framework. Instead, the drift matrix is modeled and the AR coefficients can be obtained via a nonlinear transformation of the drift matrix. Additionally, a non-linear transformation is still needed to convert IVAR to ISD. Like DSEM, CT-RDSEM cannot model IVAR (or ISD) and AR coefficients simultaneously as predictors in one step.

Our literature survey suggested that there is much research interest in studying interindividual differences in both aspects of intraindividual variability: the amplitude of fluctuation in a process variable itself (e.g., ISD or IVAR) and temporal dependency (e.g., the inertia). However, our methodological review suggested that there remains a lack of existing latent variable approaches to do so, and the prevalent method is still the

conventional approach of modeling observed IIV indicators. To our knowledge, only two prior studies have examined this issue using multilevel or latent variable approaches. Wang et al. (2012) developed a multilevel model that has both IVAR and the inertia as random effects (or latent outcomes) via a full Bayesian method. Krone et al. (2018) proposed a Bayesian dynamic model to study emotion dynamics in which IVAR and the inertia were modeled as latent outcomes. However, the performance of these methods has not been evaluated through simulation studies. Furthermore, how to model IVAR (or ISD) and the inertia as latent multiple predictors has not been previously studied. Therefore, there remains a gap in the literature regarding how to properly model both aspects of intraindividual variability, amplitude of fluctuation (e.g., ISD or IVAR) and temporal dependency (e.g., the inertia), as latent multiple predictors.

To fill in the research gap, in the current study, we aim to introduce and evaluate novel approaches to modeling the intraindividual mean, amplitude of fluctuation (e.g., ISD or IVAR), and temporal dependency (e.g., inertia) as latent multiple predictors with intensive longitudinal data. We proposed a one-step full Bayesian approach to simultaneously model the three dynamic components as latent predictors and implemented the approach in the Bayesian software program JAGS. Considering the complexity of the model in the one-step approach, we also developed two two-step hybrid-Bayesian approaches, in which a smaller model is estimated in each step, thereby reducing the computational burden in each step. Stepwise estimation procedures have been applied across various types of latent variable models and offer several advantages, such as improving model convergence, reducing misspecification, and allowing greater flexibility in handling complex model structures (e.g., Bakk & Kuha, 2018; Rosseel & Loh, 2022; Usami, 2023). An additional potential benefit of using the proposed two-step approaches is that the two steps can both be implemented in popular SEM software Mplus, allowing researchers who are not familiar with Bayesian platforms to more easily adopt the

approaches. A simulation study was conducted to compare the performance of the three novel methods to that of the conventional regression approach of using observed IIV indicators as predictors. We also included illustrative examples to showcase how the approaches can be implemented with shared code and data online.

The rest of the article is organized as follows. First, we introduce the population model of interest and demonstrate the purpose of fitting the population model. We then describe the four studied approaches: the conventional approach, the one-step full Bayesian approach, the hybrid-Bayesian approach with a single draw, and the hybrid-Bayesian approach with multiple draws. Next, we conduct a simulation study to compare the performance of the four approaches and provide modeling recommendations based on the simulation results. After that, we use an illustrative example to demonstrate the use of the four approaches. We conclude the article with discussing the implications and limitations of the study.

Methods

In the present study, we focus on examining whether intraindividual variability of variable X is predictive of an outcome variable, above and beyond the intraindividual mean of X. To study the intraindividual mean and variability of variable X, repeated measures of X can be collected over T time points from N individuals. With the longitudinal data, researchers can study interindividual differences in the intraindividual dynamic processes and how the dynamic components are associated with individuals' future outcomes. To this end, the intraindividual mean and variability (i.e., both amplitude of fluctuation and temporal dependency) of X can be extracted for each individual, then included as predictors in a multiple regression model to predict outcomes. There are multiple ways to extract intraindividual dynamic components for each individual, which will be discussed below in this section. Before that, we introduce the

population multiple regression model of interest, which is usually the main analysis model if researchers are interested in the predictivity of intraindividual dynamics.

The population regression model of interest

In the population model shown in Eq.1, intraindividual dynamic components, including the intraindividual mean (μ_i), temporal dependency (ϕ_i), and amplitude of fluctuations (ISD_i or IVAR_i), are included as predictors to predict an outcome in a multiple regression model. Note that in the population model, the predictors are true scores rather than estimates. In some methods described below in this section, however, the estimated intraindividual mean and variability instead of the true scores, are modeled as predictors.

$$\text{Outcome}_i = \beta_0 + \beta_1 * \mu_i + \beta_2 * \phi_i + \beta_3 * \text{ISD}_i \text{ (or IVAR}_i) + e_i \tag{1}$$

$$e_i \sim N(0, \sigma^2) \tag{2}$$

In Eq.1, μ_i is the true intraindividual mean which quantifies the average level of variable X over T time points for individual i. The first-order AR(1) coefficient ϕ_i is the carryover effect of observations of X at time $t-1$ on observations of X at time t, capturing the temporal dependency or inertia for individual i. ISD_i (or IVAR_i) is the true intraindividual standard deviation (or intraindividual variance) which describes the amplitude of short-term fluctuations in X for individual i. If systematic trends are identified in X, the trends should be statistically removed before modeling the intraindividual dynamics. For example, a polynomial change curve (e.g., linear) can be fitted to the time series data to remove the polynomial growth or decline trend and the differences between the raw time series data and the expected scores from the fitted curve form the detrended series.

In the IIV literature, both ISD and IVAR have been frequently used to quantify the amplitude of fluctuation. As shown in our literature survey, ISD is more frequently used as a predictor than IVAR, because ISD is in the original scale of the study variable and can have better interpretability. IVAR is more often modeled as an outcome than ISD due to its relatively tractable distribution and it has also been used as a predictor in empirical IIV studies (e.g., Brose et al., 2012; Dejonckheere et al., 2019). Therefore, in the present study, we use both ISD and IVAR as indicators of net intraindividual variability and examine the performance of different approaches of modeling them as predictors.

In the multiple regression model, the outcome of interest is regressed on μ_i, ϕ_i, and ISD_i (or IVAR_i). The intercept β_0 is the baseline of the outcome when all predictors are set to zero. The three regression coefficients, β_1, β_2, and β_3, measure the effects of the predictors on the outcome after controlling for the effects of the other two predictors. Specifically, β_1 measures how μ_i predicts the outcome above and beyond ϕ_i and ISD_i (or IVAR_i), β_2 measures how ϕ_i predicts the outcome after controlling for the effects of μ_i and ISD_i (or IVAR_i), and β_3 measures how ISD_i (or IVAR_i) predicts the outcome after accounting for μ_i and ϕ_i. The residual e_i is the deviation of the predicted outcome value to the observed value for individual i, with mean zero and standard deviation σ. In this population model, researchers are often most interested in estimating and testing the three regression coefficients (i.e., β_1, β_2, and β_3) to study whether the intraindividual mean and variability are predictive of the outcome.

Next, we introduce the conventional approach and three novel modeling approaches to estimating the population model. We start with the conventional approach in which the observed IM, AR(1) coefficient, and ISD (or IVAR) are used to predict the outcome. Next, we propose a full Bayesian method in which the true IM, AR(1) coefficient, and ISD (or IVAR) are modeled as latent variables to predict the outcome. Finally, we introduce two two-stage hybrid-Bayesian approaches in which estimates of IM, AR(1) coefficient, and ISD

(or IVAR) are obtained from a multilevel dynamic model for each individual in the first stage and then are used to predict an outcome in a regression model in the second stage.

Modeling approaches

The conventional approach. In the conventional approach, the observed or estimated IM, AR(1) coefficient, and ISD (or IVAR) are obtained for each individual before estimating the population model. More specifically, the observed IM of individual i (i.e., $\bar{x}_i$) is calculated by the average of x_{it} across T time points :

$$\bar{x}_i = (\sum_{t=1}^{T} x_{it})/T, \tag{3}$$

where x_{it} is the observed score for individual i at time t. The observed AR(1) coefficient of individual i, $\hat{\phi}_i$, is calculated by

$$\hat{\phi}_i = \sum_{t=1}^{T-1}(x_{i,t+1} - \bar{x}_i)(x_{it} - \bar{x}_i)/\sum_{t=1}^{T}(x_{it} - \bar{x}_i)^2. \tag{4}$$

The observed ISD of individual i is calculated by the standard deviation of observations over T time points:

$$\hat{\text{ISD}}_i = \sqrt{\sum_{t=1}^{T}(x_{it} - \bar{x}_i)^2/(T-1)}, \tag{5}$$

and the observed IVAR (i.e., $\hat{\text{IVAR}}_i$) is the square of $\hat{\text{ISD}}_i$.

With person-specific predictors obtained from the sample, the population model is estimated by fitting the regression model below.

$$\text{Outcome}_i = \beta_0 + \beta_1 * \bar{x}_i + \beta_2 * \hat{\phi}_i + \beta_3 * \hat{\text{ISD}}_i \text{ (or } \hat{\text{IVAR}}_i) + e_i,\ e_i \sim N(0, \sigma^2) \tag{6}$$

In this approach, the predictors are estimates calculated from the sample and the estimation can be done by a statistical inference method such as the ordinary least squares

method, maximum likelihood (ML) estimation, or Bayesian estimation. In this study, we used ML estimation to estimate the regression model under the conventional approach.

As reviewed in the introduction, the predictors are prone to the low reliability issue, especially for the intraindividual variability predictors (i.e., $\hat{\phi}_i$, $\hat{\text{ISD}}_i$, and $\hat{\text{IVAR}}_i$). Thus, we expect that the statistical inference results of the regression coefficients from this approach may be incorrect.

The One-step Full Bayesian (FB) approach. In the proposed one-step full Bayesian method, the true IM, AR(1) coefficient, and ISD (or IVAR) are modeled as latent variables via a multilevel autoregressive model (the within-person model), and the within-person dynamics and between-person relations (the between-person model) are modeled simultaneously.

The within-person model can be written as

$$x_{it} = \mu_i + \phi_i * (x_{i,t-1} - \mu_i) + \varepsilon_{it} \tag{7}$$

$$\varepsilon_{it} \sim N(0, \pi_i) \tag{8}$$

In Eq.7, μ_i is the latent or true intraindividual mean and is time-invariant, and x_{it} is latent mean centered to disaggregate within-person and between-person relations and reduce bias in estimating the AR(1) coefficients (Asparouhov et al., 2018). After latent mean centering, $x_{i,t-1} - \mu_i$ represents the temporal deviation for individual i at time $t-1$ from the intraindividual mean μ_i. The AR(1) coefficient ϕ_i describes the carryover effect of temporal deviations (i.e., $x_{i,t-1} - \mu_i$) at time $t-1$ on deviations at time t for individual i. Level-1 residual ε_{it} is the deviation of the predicted value for individual i at time t from the observed value x_{it}. ε_{it} is assumed to be normally distributed with mean zero and a person-specific variance π_i (the shock or residual variance of individual i). Note that the

intraindividual mean μ_i, AR(1) coefficient ϕ_i, and residual variance π_i are all allowed to vary across individuals.

A logarithmic transformation was applied to π_i so that the within-person residual variance π_i remains positive while $log\pi_i$ can range from negative infinity to positive infinity (Hedeker et al., 2008; Hoffman, 2007; Wang et al., 2012). Following the standard distribution specification of dynamic components in DSEM (Asparouhov et al., 2018; Hamaker et al., 2018), the three latent variables, μ_i, ϕ_i, and $log\pi_i$ are assumed to follow a multivariate normal distribution with a mean vector $\boldsymbol{\gamma}$ and a covariance matrix $\boldsymbol{\Phi}$.

Assuming stationarity, ISD_i (or IVAR_i) can be expressed as a nonlinear combination of ϕ_i and $log\pi_i$ (Wang et al., 2012):

$$\text{ISD}_i = \sqrt{\frac{exp(log\pi_i)}{1-\phi_i^2}}, \tag{9}$$

$$\text{IVAR}_i = \frac{exp(log\pi_i)}{1-\phi_i^2}. \tag{10}$$

After obtaining ISD_i (or IVAR_i), the between-person model is specified as:

$$\text{Outcome}_i = \beta_0 + \beta_1 * \mu_i + \beta_2 * \phi_i + \beta_3 * \text{ISD}_i \text{ (or IVAR}_i) + e_i,\ e_i \sim N(0, \sigma^2). \tag{11}$$

Note that the computation of ISD_i (or IVAR_i) was done in each iteration of Markov Chain Monte Carlo (MCMC) sampling. Thus, at the between-person level, all three predictors in Eq.11 are modeled as latent variables.

To estimate the model, we used non-informative priors in the present study, to allow the data "speak for themselves" without imposing strong prior beliefs. Thus, the priors have minimal influence on the posterior distributions. Alternative prior specifications (e.g., informative priors) are also available and may be appropriate depending on the context. Specifically, for the means of μ_i, ϕ_i, and $log\pi_i$ (i.e., γs), the prior was respectively a

univariate normal distribution $N(0, 10^{10})$. For the covariance matrix of μ_i, ϕ_i, and $log\pi_i$ (i.e., Φ), a non-informative inverse Wishart distribution was used with the scale matrix $\boldsymbol{I}_3$ (a 3-by-3 identity matrix) and degrees of freedom = 4 (Gelman et al., 1995; Zhang, 2021). Each regression coefficient in Eq.11, β_0 to β_3, had a normal prior $N(0, 10^{10})$. For the regression residual variance σ^2, an inverse gamma prior $IG(.001, .001)$ was specified. We used the Gibbs sampler algorithm to obtain the posterior distributions for the model parameters and latent variables (Plummer, 2017). The AR(1) coefficients were constrained to be within (-1, 1) when generating posterior samples to (1) follow the stationarity assumption (Hamilton, 2020) and (2) ensure ISD_i (or IVAR_i) was computed to be positive in each iteration. The one-step full Bayesian approach produces posterior distributions of the regression coefficients. Therefore, statistical inference can be made about the regression coefficients by examining their posterior means, medians, standard deviations, and credible intervals among other quantities. We implemented the estimation in jags using the "R2jags" package in R (Su et al., 2024).

The one-step full Bayesian approach models the within-person dynamics and between-person relations simultaneously. When both the within-person and between-person models are complex (e.g., both the within-person and between-person models involve multivariate variables), we suspect that computational problems may occur. To ease the computations, we introduce two two-step hybrid-Bayesian approaches, in which the within-person dynamics are estimated in Step 1 and the regression model of interest is estimated in Step 2, reducing the model complexity in each step.

The two-step hybrid-Bayesian approach with a single draw (HBS). In Step 1, a two-level dynamic model is fit. The within-person model specification was identical to that in the full Bayesian approach (i.e., Eqs.7 and 8). The three dynamic components, μ_i, ϕ_i, and $log\pi_i$, are assumed to follow a multivariate normal distribution with a mean vector $\boldsymbol{\gamma}$ and a covariance matrix $\boldsymbol{\Phi}$. In this step, the outcome variable of interest was added as an

auxiliary variable at the between-person level to correlate with the person-specific dynamic components (Wang et al., 2025).

$$\begin{bmatrix} \mu_i \\ \phi_i \\ log\pi_i \\ \text{Outcome}_i \end{bmatrix} = \begin{bmatrix} \gamma_\mu \\ \gamma_\phi \\ \gamma_{log\pi} \\ \gamma_O \end{bmatrix} + \begin{bmatrix} u_{1i} \\ u_{2i} \\ u_{3i} \\ u_{4i} \end{bmatrix} \tag{12}$$

$$\begin{bmatrix} u_{1i} \\ u_{2i} \\ u_{3i} \\ u_{4i} \end{bmatrix} \sim MvN(\mathbf{0}, \mathbf{\Phi}) \tag{13}$$

This model can be fit in the DSEM framework using the Bayesian estimator in M*plus*. Non-informative priors are used by default in M*plus*. For each of the four fixed effects (i.e., γs), the prior follows a univariate normal distribution $N(0, 10^{10})$. For the Level-2 covariance matrix Φ, the default prior is an inverse Wishart distribution $IW(\mathbf{0}, -p-1)$, where p is the size of the covariance matrix Φ. This prior is essentially the uniform prior ($-\infty$,∞) for all Φ parameters, resulting in a non-informative prior (Asparouhov & Muthén, 2023). With Bayesian estimation, posterior distributions of the random effects (e.g., μ_i, ϕ_i, and $log\pi_i$) can be obtained for each individual.

Note that ISD_i or IVAR_i cannot be directly modeled in the DSEM framework. But estimates of ϕ_i and $log\pi_i$ can be obtained from the posterior distributions (e.g., posterior medians) and estimates of ISD_i (or IVAR_i) can be computed based on Eq.9:

$$\hat{\text{ISD}}_i = \sqrt{\frac{exp(\hat{log\pi_i})}{1-\hat{\phi}_i^2}}, \tag{14}$$

$$\hat{\text{IVAR}}_i = \frac{exp(\hat{log\pi_i})}{1-\hat{\phi}_i^2}, \tag{15}$$

under the stationarity assumption. For individuals whose AR(1) coefficient posterior medians are ≥ 1 or ≤ -1, the time series would be nonstationary and the ISD_i (or IVAR_i) estimates would be negative or undefined. Therefore, these individuals were excluded from the step-2 analysis.

In Step 2, posterior medians of μ_i and ϕ_i (i.e., $\hat{\mu}_i$ and $\hat{\phi}_i$) and the calculated $\hat{\text{ISD}}_i$ (or $\hat{\text{IVAR}}_i$) were modeled as predictors to predict the outcome variable:

$$\text{Outcome}_i = \beta_0 + \beta_1 * \hat{\mu}_i + \beta_2 * \hat{\phi}_i + \beta_3 * \hat{\text{ISD}}_i \text{ (or } \hat{\text{IVAR}}_i) + e_i,\ e_i \sim N(0, \sigma^2) \tag{16}$$

Note that in Eq.16, the three predictors are estimates obtained from DSEM, not latent variables.

We used M*plus* to fit the multiple regression model with maximum likelihood estimation. Although it is easy to implement, this approach ignores the uncertainty in the single estimates (i.e., posterior medians), which may underestimate uncertainty measures (e.g., standard errors) thus leading to poor coverage issues (Rubin, 1987; Schafer & Graham, 2002). To circumvent this issue, we proposed a two-step approach with multiple draws.

The two-step hybrid-Bayesian approach with multiple draws (HBM). In Step 1, we obtain the posterior distributions of the three dynamic components by fitting the same model in the first step of the single-draw approach (i.e., Eqs.7-8 and Eqs.12-13). Different from the HBS approach, multiple (e.g., $M = 50$) values of ϕ_i and $log\pi_i$ were drawn from their posterior distributions. Then, multiple ISD_i (or IVAR_i) estimates were calculated for

each individual accordingly, yielding M datasets:

$$\hat{\text{ISD}}_{im} = \sqrt{\frac{exp(\hat{log\pi_{im}})}{1-\hat{\phi}_{im}^2}}, \tag{17}$$

$$\hat{\text{IVAR}}_{im} = \frac{exp(\hat{log\pi_{im}})}{1-\hat{\phi}_{im}^2}, \tag{18}$$

where $m = 1, 2, \ldots, M$. Similar to the HBS approach, individuals whose AR(1) draws have absolute values larger than or equal to one were excluded from further analysis. Specifically, individuals with any out-of-range AR(1) draws were excluded from all M datasets because the multiple imputation function in M*plus* requires an equal number of observations across all datasets.

In Step 2, a multiple regression model was fit to each dataset:

$$\text{Outcome}_{im} = \beta_{0m} + \beta_{1m} * \hat{\mu_{im}} + \beta_{2m} * \hat{\phi_{im}} + \beta_{3m} * \hat{\text{ISD}}_{im} \text{ (or } \hat{\text{IVAR}}_{im}) + e_{im},\ e_{im} \sim N(0, \sigma_m^2). \tag{19}$$

Then the regression results from M datasets were aggregated using the standard pooling method for multiple imputation (Rubin, 1996), with the point estimate

$$\bar{Q} = \sum_{m=1}^{M} \hat{Q}_m / M \tag{20}$$

and the standard error estimate

$$\sqrt{\sum_{m=1}^{M} \hat{U}_m / M + \frac{M+1}{M} * \sum_{m=1}^{M} (\hat{Q}_m - \bar{Q})^2 / (M-1)}, \tag{21}$$

where $Q \in \{\beta_1, \beta_2, \beta_3\}$ and $U \in \{s_{\beta_1}^2, s_{\beta_2}^2, s_{\beta_3}^2\}$.

The first step was also implemented in the DSEM framework in M*plus* with the Bayesian estimation. In the second step, we used the multiple imputation function in

M*plus* to fit the multiple regression model of interest with maximum likelihood estimation and aggregate the results.

The steps of the full Bayesian and hybrid-Bayesian modeling approaches are summarized in Table 1.

A Simulation Study

We conducted a simulation study to compare the performance of the four modeling approaches (i.e., the conventional approach, the one-step full Bayesian approach, the hybrid-Bayesian approach with a single draw, and the hybrid-Bayesian approach with multiple draws) under different data conditions. With the results, modeling recommendations will be provided. In this section, we first introduce the simulation study design and evaluation criteria. Then, we describe the simulation data generation and show the simulation results.

Simulation Design

We employed a factorial experimental design with four factors: the sample size ($N = 50, 100, 200, 300, 500$), number of time points ($T = 10, 20, 30, 50, 100$), variability indicator (i.e., ISD or IVAR for the net variability), and effect size of the three regression coefficients (zero or small-medium $f^2 = .05$ to .1; Cohen, 1988; Maxwell, 2000). Across the three regression coefficients, there are $2 \times 2 \times 2 = 8$ effect size combinations. Therefore, a total of 400 ($5 \times 5 \times 2 \times 8$) conditions were evaluated. Each condition had 500 replications.

For the full Bayesian or hybrid-Bayesian modeling approaches, two chains were used for MCMC sampling. For each chain, the number of iterations was set to 10,000 and the thinning period was set to 10. The first half draws were burnt to stabilize the results. Hence, a total of 1,000 posterior draws were obtained for each parameter from the two chains. For the hybrid-Bayesian approach with multiple draws, we set the number of draws

$M = 50$, based on the recommendations in the multiple imputation literature (Rubin, 1987, 1996).

In this simulation study, we were interested in whether the three regression coefficients (i.e., $\beta_1, \beta_2, \beta_3$) can be recovered well using the four modeling approaches. We used the relative bias and 95% confidence or credible interval (CI) coverage rate as the evaluation criteria for parameters with nonzero true value. For parameters with zero true values, the empirical bias and empirical Type I error rate were used as the evaluation criteria.

Specifically, the relative bias measures the extent to which parameter estimates deviate from the true values. It is calculated by subtracting the true value from the average estimate (the average of the point estimates across all replications within a condition) and then dividing this difference by the true value for each parameter. Relative biases falling within the range of [-10%, 10%] were considered ignorable biases (Muthén & Muthén, 2002). When the true parameter value is zero, the relative bias cannot be obtained because the denominator is zero. In this case, we used the empirical bias to measure the accuracy of the parameter estimates, which is calculated by the difference between the average estimate across all replications of a condition and the true value.

The 95% CI coverage rate and the empirical Type I error rate measure how well statistical inferences are made for a parameter. We used the 95% CI coverage rate, the proportion of the 95% confidence or credible intervals that contain the true value for each parameter within a condition to evaluate the inference performance for parameters with nonzero true values. Good 95% CIs are expected to yield a coverage probability close to 95%. We treated coverage rates falling within [91%, 98%] as satisfactory (Muthén & Muthén, 2002). For parameters with zero true values, the empirical Type I error rate, the percentage of 95% CIs that do not contain zero, was used. Empirical Type I error rates within [2.5%, 7.5%] were treated as acceptable for the nominal significance level of 5%

(Bradley, 1978).

Data Generation

We generated data based on Eqs.7-11. First, person-specific μ_i, ϕ_i, and $log\pi_i$ were generated from a multivariate normal distribution. The true means and between-person variances of the person-specific dynamic components were adopted from the real data analysis results reported in Table 1 of Hamaker et al. (2018). The true mean vector of the three person-specific components is $\boldsymbol{\gamma} = [1, .3, -2]$, and the true variances (i.e., diagonal elements of Φ) are $[.6, .04, 1.9]$. The population correlations among the dynamic components were set to .3 (medium size). We constrained the generated values of ϕ_i to be within $(-1, .95)$ to (1) follow the stationarity assumption (Hamilton, 2020) and (2) avoid potential computational issues arising by ϕ_i being too close to one (Asparouhov & Muthén, 2024b), such as avoiding having close to 0 denominator values when computing ISD_i (or IVAR_i) in the Bayesian approaches. AR(1) parameters that were generated outside $(-1, .95)$ were discarded (about 0.05% of the generated values) and regenerated. There were about 6.7% negative autocorrelations simulated, within the range (6.5% to 17.4%) reported in Hamaker et al. (2018).

After obtaining the three person-specific dynamic components, T occasions of e_{it} was generated by Eq.8. There are several ways to generate stationary time series data. In this simulation, we used the method in Fang & Wang (2024). Specifically, we set the starting values (i.e., $x_{i,t=-50}$) to be 0 and generated the next $49 + T$ occasions of data for x_{it} based on Eq.7 for individual i. Then, the first 50 observations of x_{it} were discarded to ensure the remaining T observations are stationary for individual i.

In the final step, we computed the true ISD_i (or IVAR_i) based on Eq.9 and generated the outcome variable based on Eqs.1 and 2. We set the regression residual variance $\sigma^2 = 1$ and the true values of $\beta_1, \beta_2, \beta_3$ were determined by their effect sizes in the

simulation condition.

Simulation Results

As mentioned earlier in the paper, the hybrid-Bayesian approaches remove out-of-range AR(1) coefficient draws in the second step. Across the simulation conditions, for the HBS approach, less than 0.01% of the AR(1) coefficient posterior medians had absolute values larger than or equal to one. In the HBM approach, on average, about 1% individuals had out-of-range AR(1) coefficient draws and were excluded from subsequent regression analysis.

Using ISD as a predictor. Figure 1 displays the relative biases and 95% CI coverage rates from the conditions where all βs have small-medium effect sizes and ISD is a predictor, and Figure 2 shows the empirical biases and empirical Type I error rates from the conditions where all βs have true values of zero and ISD is a predictor.

As seen from Figure 1, the conventional and HBS approaches failed to recover the parameters. For example, the conventional approach had non-ignorable biases and low coverage rates when recovering the regression coefficient of the AR(1) coefficient predictor (i.e., β_2). The relative bias was -18.6% and 95% CI coverage rate was 80.8% with $T = 100$ and $N = 500$. The coverage rate increased with larger T but decreased with larger N. The HBS approach had a 30.0% relative bias and a 62.1% coverage rate of β_2 when $T = 100$ and $N = 500$. Similar to the conventional approach, coverage rates of the HBS approach increased with larger T but decreased with larger N. Moreover, as seen in Figure 2, the HBS approach had inflated Type I error rates for β_2 (e.g., 8.8% when $T = 100$ and $N = 500$).

Compared to the conventional and HBS approaches, the HBM approach had better performance in recovering the regression coefficients. Specifically, the HBM approach had ignorable relative biases and satisfactory coverage rates for all three regression coefficients

when $T \geq 50$. In terms of the empirical Type I error rates, the HBM approach did not have inflated Type I error rates for β_2 and β_3, although this approach tended to be conservative with fewer occasions of data (e.g., $T = 10$).

The one-step FB approach generally worked well in recovering all regression coefficients with sufficient N and T. The relative biases were ignorable and coverage rates were satisfactory when $T \geq 30$ and $N \geq 100$, except for only a few conditions slightly beyond the range of [-10%, 10%]. As shown in Figure 2, the FB approach had well-controlled Type I error rates when $T \geq 10$ and $N \geq 100$.

In sum, the results suggested that the conventional approach and the HBS approach had problems in recovering the regression coefficients, especially for the regression coefficient of the AR(1) coefficient predictor (i.e., β_2), and thus are not recommended. The HBM approach successfully recovered all regression coefficients when $T \geq 50$. The one-step FB approach generally worked well and its data requirement (i.e., $T \geq 30$ and $N \geq 100$) was lower than the HBM approach.

In other effect size conditions, such as one or two βs having true values of zero, the conventional and HBS approaches consistently had unsatisfactory performance, especially in recovering β_2. In contrast, the HBM and FB approaches recovered all regression coefficients well, despite slightly different data requirements. The relative biases, empirical biases, 95% CI coverage rates, and empirical Type I error rates from these effect size conditions can be found in the online supplemental material (Appendix B).

Using IVAR as a predictor. The results from conditions where all βs have small-medium effect sizes and IVAR is a predictor are displayed in Figure 3 and those from conditions where all βs have zero true values and IVAR is a predictor are displayed in Figure 4. The overall findings from modeling IVAR as a predictor were similar to those from modeling ISD as a predictor, except for differences in the data requirements on N and T.

First, the conventional approach failed to recover β_2 when all βs had small-medium effect sizes. The relative bias was -17.8% and the coverage rate was 76.2% when $T = 100$ and $N = 500$. Similarly, the HBS approach failed to recover β_2. The relative bias was -25.3% and the coverage rate was 65.0% when $T = 100$ and $N = 500$. The coverage rates of the two approaches increased with a larger T but decreased with a larger N. In terms of the empirical Type I error rates, the conventional approach had well-controlled Type I error rates, whereas the HBS approach had inflated Type I error rates, especially for β_2.

The HBM and FB approaches had satisfactory performance with sufficient N and T. When all βs had small-medium effect sizes, the HBM approach recovered all βs well with satisfactory coverage rates and well-controlled Type I error rates except for a few lower than 2.5% empirical Type I error rates when T is small (e.g., 10 or 20). The relative biases for β_1 and β_2 were generally ignorable when $T \geq 30$; but the relative biases for β_3 approached the ignorable range when $T \geq 100$ (-14% to -4% relative biases when $T = 100$ and $N \geq 100$) . The FB approach worked well when $T \geq 30$ and $N \geq 200$, with ignorable relative biases, satisfactory coverage rates, and well-controlled Type I error rates.

In sum, when IVAR is a predictor, the conventional and HBS approaches had problems in recovering the regression coefficients, especially the one of the AR(1) predictor (i.e., β_2), and thus are not recommended. The HBM approach can successfully recover all regression coefficients but the number of time points required can be 100 or more. The one-step full Bayesian approach generally worked well and its data requirement (i.e., $T \geq 30$ and $N \geq 200$) was lower than that of the HBM approach.

Illustrative examples

To illustrate the use of the four modeling approaches and numerically compare the results from the approaches, we used simulated data based on real data analysis results from Ong et al. (2025). We presented two illustrative examples: one used ISD as a

predictor, and the other used IVAR. In Ong et al. (2025), negative affect was measured on 56 consecutive days in a sample of 799 adults. At the within-person level, daily negative affect was modeled by a multilevel AR(1) model to capture the temporal dependency across the 56 days. At the between-person level, three dynamic components, including the intraindividual mean (i.e., IM), inertia (i.e., the first-order AR coefficient), and variability (i.e., ISD) of daily negative affect, were used to predict depressive symptoms measured after the 56-day window. Results found that higher intraindividual means and greater inertia of negative effect were associated with higher scores of depression. However, variability measured by ISD was not significantly linked to depressive symptoms after accounting for the intraindividual mean and inertia.

Two datasets were generated: one for modeling ISD as a predictor, and the other for modeling IVAR as a predictor. The true values used to generate the data in the illustrative examples were based on the results in Ong et al. (2025). Specifically, μ_i, ϕ_i, and $log\pi_i$ in Eq.12 were generated from a multivariate normal distribution. The means of the dynamic components were set to be 1.25, .3, and -3.2, respectively, and the variances were set to be .16, .06, 2.1, respectively. The variance of μ_i was modified from the real data analysis result .04 to .16 to increase the between-person variance of the intraindividual means. The correlations among the three dynamic components were .55 between μ_i and ϕ_i, .81 between μ_i and $log\pi_i$, and .40 between ϕ_i and $log\pi_i$, respectively. Then, the true ISD_i (or IVAR_i) was computed using Eq.9 (or Eq.10). The regression coefficients of IM, AR(1), and ISD (or IVAR) in the multiple regression models were set to be 10, 8, and -2 respectively. The effect sizes of the regression coefficients ($f^2 = .10$, .04, .002 in Example 1 where ISD was modeled as a predictor; $f^2 = .14$, .04, .03 in Example 2 where IVAR was modeled as a predictor) correspond to the medium ($\hat{f^2}_{IM} = .16$) and small ($\hat{f^2}_{AR1} = .02$, $\hat{f^2}_{ISD} = .004$) effect sizes in Ong et al. (2025). The intercept and residual variance of the multiple regression models in both examples were set to 0 and 60, respectively. Using those values,

negative affect scores and depressive symptom scores in both examples were generated following the same procedure used in the simulation study. Each simulated data had 800 individuals and 56 time points, similar to Ong et al. (2025).

In the modeling approaches that involve MCMC sampling, the numbers of iterations, burn-in period, and thinning period for each of the two chains were set to 10,000, 5,000, and 10, respectively, consistent with the simulation study. In the HBS approach, the number of posterior samples that were used to get the posterior medians was 1000. In the HBM approach, the number of multiple draws was 50. [1]

Table 2 shows the results from the four modeling approaches for the two examples. First, the conventional approach failed to recover β_3 regardless of whether ISD or IVAR was a predictor. Second, the HBS approach failed to recover β_2 when IVAR was a predictor, with a substantial relative bias (50%) and a 95% CI not covering the true β_2. When ISD was a predictor, the HBS approach also overestimated β_2, resulting in a 40% relative bias in the parameter estimate. In contrast, the HBM and FB approaches recovered all three parameters well with their 95% CIs covering the true regression coefficient values. These two illustration examples showed that when inappropriate methods were used to model intraindividual variability as predictors, biased parameter estimates or inaccurate inferences could be obtained.

Discussion

The predictivity of intraindividual variability on health or psychological outcomes is widely studied in many subfields of psychology, such as affect dynamics, cognitive performance, and aging research. However, methodological research on modeling intraindividual variability remains limited. In particular, the amplitude of fluctuation quantified by the intraindividual standard deviation (or variance) and the temporal dependency (e.g., the inertia), can be challenging to be simultaneously modeled as

multiple predictors in popular latent variable frameworks, such as DSEM. To address the research gap, we introduced and evaluated three novel modeling methods, including a one-step full Bayesian approach and two two-step hybrid-Bayesian approaches, to model intraindividual variabilty as predictors.

Primary findings and implications

Our simulation study showed that the conventional regression approach of modeling observed IIV indicators as predictors led to inaccurate inferences of the regression coefficients. In the literature, researchers have noted that the observed IIV indicators can have low reliability estimates when the number of time points is not sufficient (e.g., Du & Wang, 2018; Wang & Grimm, 2012) and improper inferences can be made for regression coefficients when predictors have the low reliability issue (McDonald, 2013). Thus, our simulation results on the performance of the conventional approach were consistent with the literature.

The two hybrid-Bayesian approaches have advantages compared to the one-step full Bayesian approach. First, each step of the hybrid-Bayesian approaches has a smaller model than the one-step full Bayesian approach, leading to fewer potential computational issues. Second, in the current study, nonlinear transformations were made on the dynamic component estimates to obtain estimates of intraindividual standard deviation (ISD) and intraindividual variance (IVAR) in Step 1 and thus allow the direct modeling of ISD or IVAR instead of the shock or residual variance in Step 2. In other studies, other linear or nonlinear transformations can be made on the step-1 estimates to allow the estimation and test of other models of interest (e.g., mediation or moderation analysis) in Step 2. Between the two hybrid-Bayesian approaches, the performance of the HBS approach was not satisfactory, as it ignored the uncertainty in the single draw estimates. However, the performance of the HBM approach was satisfactory with sufficient sample sizes and

numbers of time points.

The HBM approach was implemented in M*plus* under the DSEM framework, which can enjoy the advantages of DSEM, such as handling unequal time intervals and missing data (Asparouhov et al., 2018; Fang & Wang, 2024; Hamaker et al., 2018). The two-step nature of this method makes it a feasible alternative to the one-step full Bayesian approach.

The one-step full Bayesian approach had satisfactory performance when modeling either ISD or IVAR as a predictor and the data requirements were lower than those of the HBM approach. We implemented the FB approach in R JAGS. However, it can also be implemented in other Bayesian software programs such as BUGS or Stan. Researchers who are more familiar with other Bayesian software programs can modify the JAGS code and implement the FB approach elsewhere.

Limitations and future directions

In the current study, we assumed the process variable from which the IIV indicators are calculated from has zero measurement error. In psychological or behavioral research, variables of interest often are measured with a measurement scale and thus measurement errors can exist. Wang & Liu (2025) showed that when measurement errors exist, the conventional regression approach performed even worse than when there is no measurement error. For example, they found that regular regression yielded worse performance about the coefficient of ISD when the number of time points increases under the presence of measurement errors. Our proposed approaches including the HBM and FB approaches, which model both the amplitude of fluctuation and temporal dependency, can be extended to account for measurement errors in future research.

In our hybrid-Bayesian approaches, we included step-2 variables in Step 1 modeling as auxiliary variables to model the between-person covariance matrix. In the missing data literature, the inclusion of auxiliary variables has been discussed intensively (e.g., Collins

et al., 2001). In the mixture modeling literature, whether and how to include the distal outcomes, predictors, and covariates into the latent class detection process has also been discussed (e.g., Asparouhov & Muthén, 2014). Preliminary simulation results showed that if the step-2 variables are not included in step-1 analysis, inaccurate inferences were made from the hybrid-Bayesian approaches (Wang et al., 2025). Whether and how to include step-2 variable in Step 1 modeling should be further evaluated for dynamic modeling in future research.

In the simulation study, we adopted the between-person variance parameter values of the dynamic components from a real data analysis Hamaker et al. (2018). We did not vary the sizes of the variances of the dynamic components due to the already large number of studied simulation conditions (i.e., 400). Thus, we did not examine its impact on the performance of the modeling approaches in the present study. Studying this impact could be an interesting future research direction.

Conclusions

In this article, we proposed and evaluated novel modeling approaches to simultaneously model two aspects of intraindividual variability (i.e., the amplitude of fluctuation and temporal dependency) as predictors. Our simulation results showed the conventional regression approach and the hybrid-Bayesian approach with a single draw can lead to inaccurate inference results (biased estimates and poor coverage rates) about the predictivity of intraindividual variability indicators. In contrast, the proposed hybrid-Bayesian approach with multiple draws and the proposed one-step full Bayesian approach yielded accurate inferences when the sample size and the number of time points are sufficient. We provided simulated data examples to illustrate the use of the approaches with sharing the data and model code online. We hope this study offers useful tools for researchers interested in intraindividual variability, especially in understanding how

interindividual differences in intraindividual variability relate to outcomes.

References

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 359–388.

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 329–341.

Asparouhov, T., & Muthén, B. (2024a). *Continuous time dynamic structural equation models.* Retrieved from `https://www.statmodel.com/download/CTRDSEM.pdf`

Asparouhov, T., & Muthén, B. (2024b). *Practical aspects of dynamic structural equation models.* Retrieved from `http://www. statmodel.com/download/PDSEM.pdf`

Asparouhov, T., & Muthén, B. O. (2023). *Bayesian analysis using Mplus: Technical implementation.* Retrieved from `https://www.statmodel.com/download/Bayes2.pdf`

Bakk, Z., & Kuha, J. (2018). Two-step estimation of models between latent classes and external variables. *Psychometrika*, *83*(4), 871–892.

Bielak, A. A., Hultsch, D. F., Strauss, E., MacDonald, S. W., & Hunter, M. A. (2010). Intraindividual variability is related to cognitive change in older adults: Evidence for within-person coupling. *Psychology and Aging*, *25*(3), 575-586.

Boker, S. M., Molenaar, P., & Nesselroade, J. R. (2009). Issues in intraindividual variability: Individual differences in equilibria and dynamics over multiple time scales. *Psychology and Aging*, *24*(4), 858-862.

Bradley, J. V. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology*, *31*(2), 144–152.

Brockwell, P. J., & Davis, R. A. (2002). *Introduction to time series and forecasting.* Springer.

Brose, A., Schmiedek, F., Lövdén, M., & Lindenberger, U. (2012). Daily variability in working memory is coupled with negative affect: The role of attention and motivation. *Emotion*, *12*(3), 605–617.

Carstensen, L. L., Turan, B., Scheibe, S., Ram, N., Ersner-Hershfield, H., Samanez-Larkin, G. R., ... Nesselroade, J. R. (2011). Emotional experience improves with age: evidence based on over 10 years of experience sampling. *Psychology and Aging*, *26*(1), 21-33.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. ed.). Hillsdale, NJ: Erlbaum.

Collins, L. M., Schafer, J. L., & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, *6*(4), 330-351.

Dejonckheere, E., Mestdagh, M., Houben, M., Rutten, I., Sels, L., Kuppens, P., & Tuerlinckx, F. (2019). Complex affect dynamics add limited information to the prediction of psychological well-being. *Nature Human Behaviour*, *3*(5), 478–491.

Du, H., & Wang, L. (2018). Reliabilities of intraindividual variability indicators with autocorrelated longitudinal data: Implications for longitudinal study designs. *Multivariate Behavioral Research*, *53*(4), 502–520.

Dzubur, E., Ponnada, A., Nordgren, R., Yang, C.-H., Intille, S., Dunton, G., & Hedeker, D. (2020). Mixwild: A program for examining the effects of variance and slope of time-varying variables in intensive longitudinal data. *Behavior Research Methods*, *52*(4), 1403–1427.

Estabrook, R., Grimm, K. J., & Bowles, R. P. (2012). A monte carlo simulation study of the reliability of intraindividual variability. *Psychology and Aging*, *27*(3), 560-576.

Fang, Y., & Wang, L. (2024). Dynamic structural equation models with missing data: Data requirements on N and T. *Structural Equation Modeling: A Multidisciplinary Journal*, *31*(5), 891–908.

Feng, Y., & Hancock, G. R. (2024). A structural equation modeling approach for modeling variability as a latent variable. *Psychological Methods*, *29*(2), 262-286.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian data analysis*. Chapman and Hall/CRC.

Gerstorf, D., Siedlecki, K. L., Tucker-Drob, E. M., & Salthouse, T. A. (2009). Within-person variability in state anxiety across adulthood: Magnitude and associations with between-person characteristics. *International Journal of Behavioral Development*, *33*(1), 55–64.

Hamaker, E. L., Asparouhov, T., Brose, A., Schmiedek, F., & Muthén, B. (2018). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the cogito study. *Multivariate Behavioral Research*, *53*(6), 820–841.

Hamaker, E. L., Asparouhov, T., & Muthén, B. (2021). Dynamic structural equation modeling as a combination of time series modeling, multilevel modeling, and structural equation modeling. In R. H. Hoyle (Ed.), *The handbook of structural equation modeling* (2. ed., chap. 31). Guilford Press.

Hamilton, J. D. (2020). *Time series analysis*. Princeton University Press.

Hedeker, D., Mermelstein, R. J., & Demirtas, H. (2008). An application of a mixed-effects location scale model for analysis of ecological momentary assessment (EMA) data. *Biometrics*, *64*(2), 627–634.

Hoffman, L. (2007). Multilevel models for examining individual differences in within-person variation and covariation over time. *Multivariate Behavioral Research*, *42*(4), 609–629.

Jahng, S., Wood, P. K., & Trull, T. J. (2008). Analysis of affective instability in ecological momentary assessment: Indices using successive difference and group comparison via multilevel modeling. *Psychological Methods*, *13*(4), 354-375.

Krone, T., Albers, C. J., Kuppens, P., & Timmerman, M. E. (2018). A multivariate statistical model for emotion dynamics. *Emotion*, *18*(5), 739-754.

Kuppens, P., Sheeber, L. B., Yap, M. B., Whittle, S., Simmons, J. G., & Allen, N. B. (2012). Emotional inertia prospectively predicts the onset of depressive disorder in adolescence. *Emotion*, *12*(2), 283–289.

Laurenceau, J.-P., & Bolger, N. (2005). Using diary methods to study marital and family processes. *Journal of Family Psychology*, *19*(1), 86-97.

Maxwell, S. E. (2000). Sample size and multiple regression analysis. *Psychological Methods*, *5*(4), 434-458.

McDonald, R. P. (2013). *Test theory: A unified treatment*. Psychology Press.

Muthén, L. K., & Muthén, B. O. (2002). How to use a monte carlo study to decide on sample size and determine power. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(4), 599–620.

Nesselroade, J. R. (1991). The warp and woof of the developmental fabric. In R. Downs, L. Liben, & D. Palermo (Eds.), *Visions of development, the environment, and aesthetics: The legacy of joachim f. wohlwill.* Hillsdale, NJ: Erlbaum.

Nesselroade, J. R. (2001). Intraindividual variability in development within and between individuals. *European Psychologist*, *6*(3), 187–193.

Nesselroade, J. R., & Molenaar, P. C. (2010). Emphasizing intraindividual variability in the study of development over the life span: Concepts and issues. In W. F. Overton & R. M. Lerner (Eds.), *The handbook of life-span development.* John Wiley & Sons, Inc.

Ong, A., Wang, L., Fang, Y., Zhang, G., Joiner, R., Wilcox, K., & Bergeman, C. (2025). Affect dynamics and depressive symptomatology: revisiting the inertia-instability paradox. *Emotion*, *25*(4), 1060–1064.

Plummer, M. (2017). Jags version 4.3.0 user manual [Computer software manual].

Ram, N., & Gerstorf, D. (2009). Time-structured and net intraindividual variability: Tools for examining the development of dynamic characteristics and processes. *Psychology and Aging*, *24*(4), 778-791.

Rast, P., Martin, S. R., Liu, S., & Williams, D. R. (2022). A new frontier for studying within-person variability: Bayesian multivariate generalized autoregressive conditional heteroskedasticity models. *Psychological Methods*, *27*(5), 856-873.

Rosseel, Y., & Loh, W. W. (2022). A structural after measurement approach to structural equation modeling. *Psychological Methods*, *29*(3), 561–588.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys.* New York: Wiley.

Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association*, *91*(434), 473–489.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147-177.

Su, Y.-S., Yajima, M., & Baio, G. (2024). R2JAGS: Using R to run JAGS [Computer software manual].

Usami, S. (2023). Within-person variability score-based causal inference: A two-step estimation for joint effects of time-varying treatments. *Psychometrika*, *88*(4), 1466–1494.

Wang, L., Fang, Y., & Bergeman, C. (2025). Dynamic modeling with intensive longitudinal data: One-step and two-step DSEM approaches. *Manuscript submitted for publication*.

Wang, L., & Grimm, K. J. (2012). Investigating reliabilities of intraindividual variability indicators. *Multivariate Behavioral Research*, *47*(5), 771–802.

Wang, L., Hamaker, E., & Bergeman, C. (2012). Investigating inter-individual differences in short-term intra-individual variability. *Psychological Methods*, *17*(4), 567-581.

Wang, L., & Liu, X. (2025). Modeling intraindividual variability as a predictor with longitudinal data. *Psychological Methods*.

Wright, A. J., & Jackson, J. J. (2025). Longitudinal within-person variability around personality trajectories. *Journal of Personality and Social Psychology*, *128*(1), 211–242.

Zhang, Z. (2021). A note on Wishart and inverse Wishart priors for covariance matrix. *Journal of Behavioral Data Science*, *1*(2), 119–126.

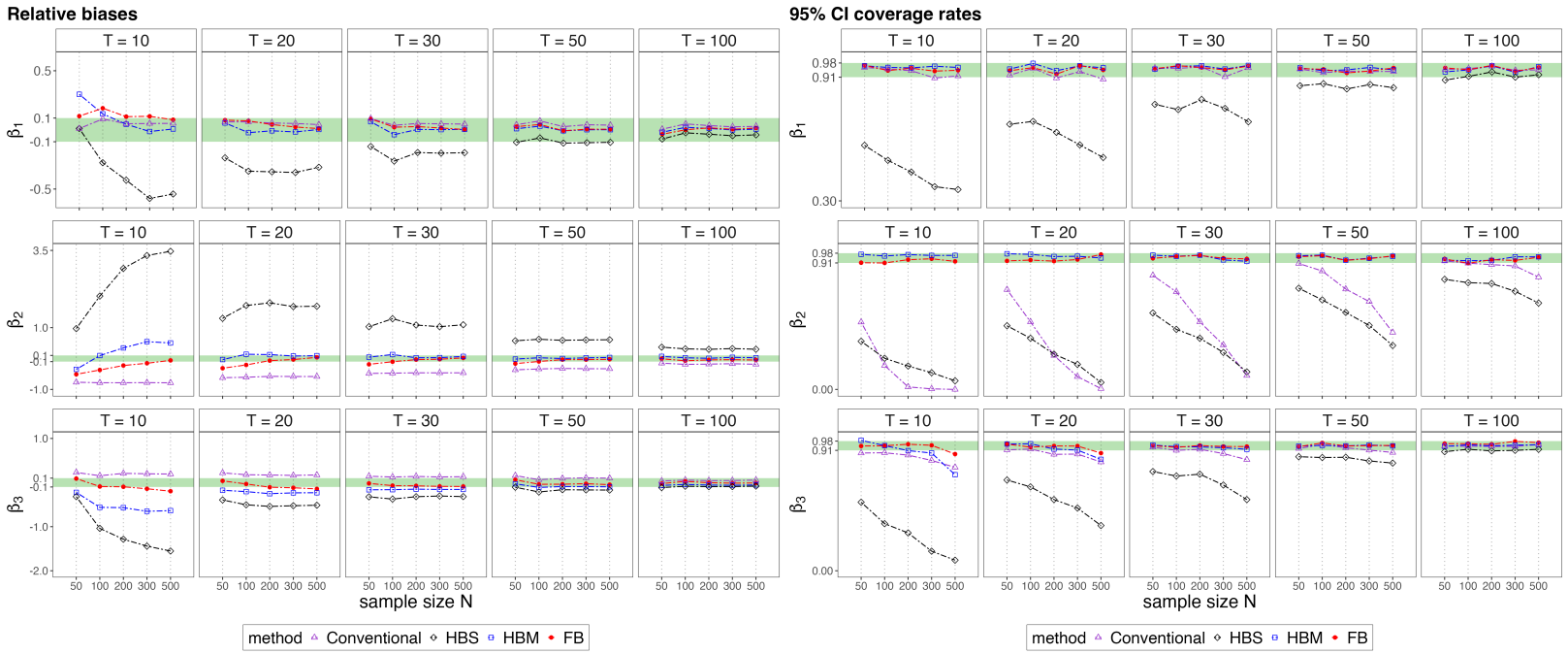
Footnotes

[1]The example data and implementation code for the four methods can be found in https://github.com/peggywangnd/IIV_DSEM.

Table 1
Steps in the full and hybrid-Bayesian modeling approaches

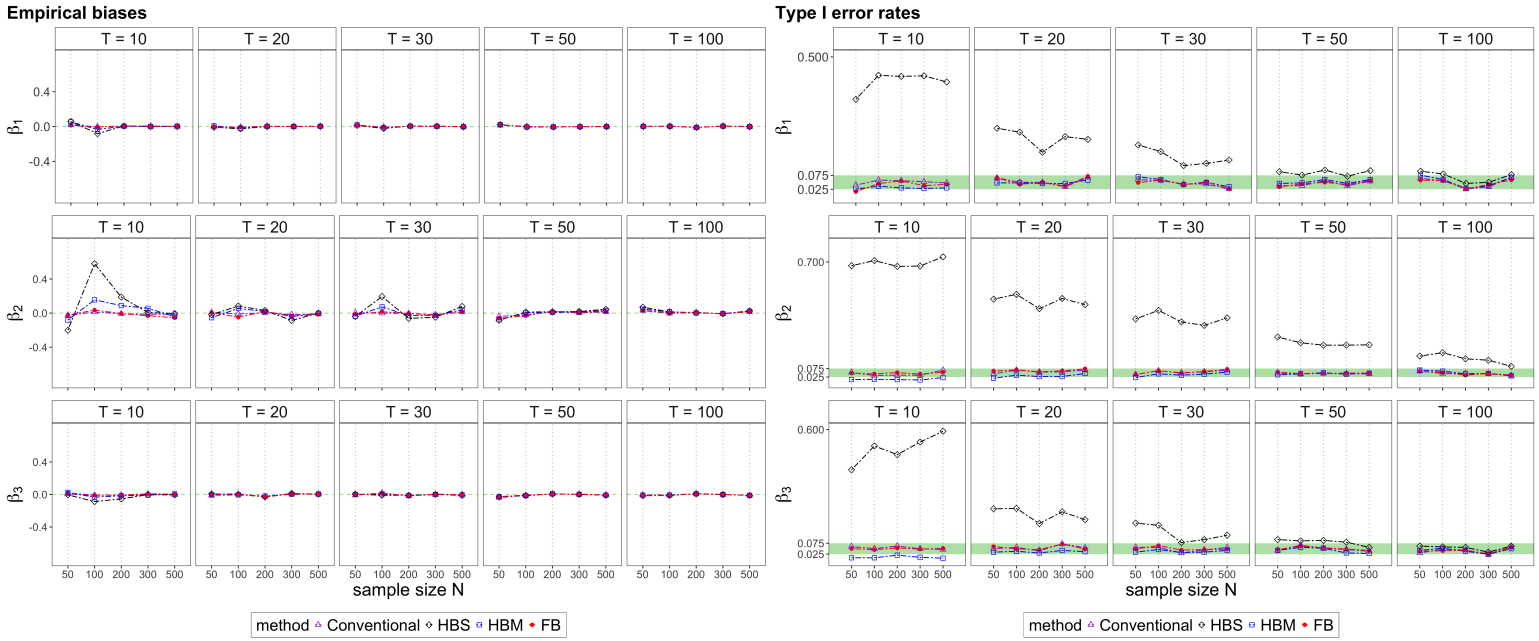
One-Step Full-Bayesian (FB)	Hybrid-Bayesian Modeling with a Single Draw (HBS)	Hybrid-Bayesian Modeling with a Multiple Draws (HBM)
A single step: **Jointly model** the intraindividual dynamic components and interindividual relations by estimating the within-person dynamic relations and between-person regression model simultaneously.	Step 1a: Fit a DSEM and obtain posterior distributions of μ_i, ϕ_i, and $log\pi_i$.	Step 1a: Fit a DSEM and obtain posterior distributions of μ_i, ϕ_i, and $log\pi_i$.
	Step 1b: $\hat{\text{ISD}}_i$ (or $\hat{\text{IVAR}}_i$) is computed based on **posterior medians** of ϕ_i and $log\pi_i$: $\hat{\text{ISD}}_i = \sqrt{\frac{exp(\hat{log\pi_i})}{1-\hat{\phi}_i^2}}, \hat{\text{IVAR}}_i = \frac{exp(\hat{log\pi_i})}{1-\hat{\phi}_i^2}$.	Step 1b: Multiple $\hat{\text{ISD}}_i$ (or $\hat{\text{IVAR}}_i$) are computed based on **multiple posterior draws** of ϕ_i and $log\pi_i$, yielding multiple datasets $\hat{\text{ISD}}_{im} = \sqrt{\frac{exp(\hat{log\pi_{im}})}{1-\hat{\phi}_{im}^2}}$, $\hat{\text{IVAR}}_{im} = \frac{exp(\hat{log\pi_{im}})}{1-\hat{\phi}_{im}^2}$, $m = 1, 2, \ldots, M$.
ϕ_i and ISD_i (or IVAR_i) are calculated within each iteration of MCMC sampling and are modeled as latent variables.	Step 2: Fit the regression model of interest with $\hat{\mu}_i$, $\hat{\phi}_i$, and $\hat{\text{ISD}}_i$ (or $\hat{\text{IVAR}}_i$) as predictors.	Step 2: Fit the regression model with $\hat{\mu}_i$, $\hat{\phi}_i$, and $\hat{\text{ISD}}_i$ (or $\hat{\text{IVAR}}_i$) as predictors for each dataset, and aggregate the results using the standard pooling method in Rubin (1996).

Figure 1. Relative biases and 95% CI coverage rates when all regression coefficients had small-medium effect sizes and ISD is a predictor.



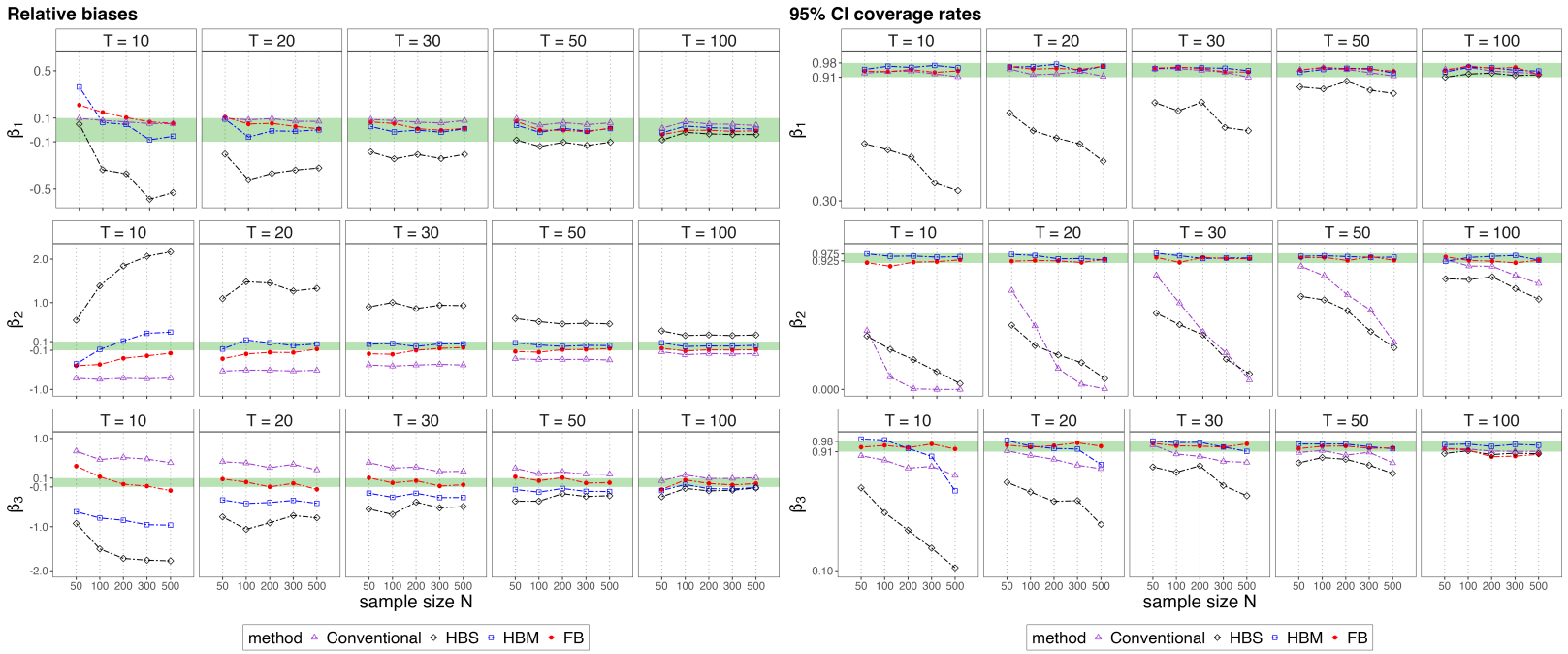
Note. The shaded area marks the boundaries of $[-.10, .10]$ for relative biases and $[.91, .98]$ for 95% CI coverage rates. Conventional = the conventional approach; HBS = the hybrid-Bayesian approach with a single draw; HBM = the hybrid-Bayesian approach with multiple draws; FB = the one-step full Bayesian approach.

Figure 2. Empirical biases and empirical Type I error rates when all regression coefficients were zero and ISD is a predictor



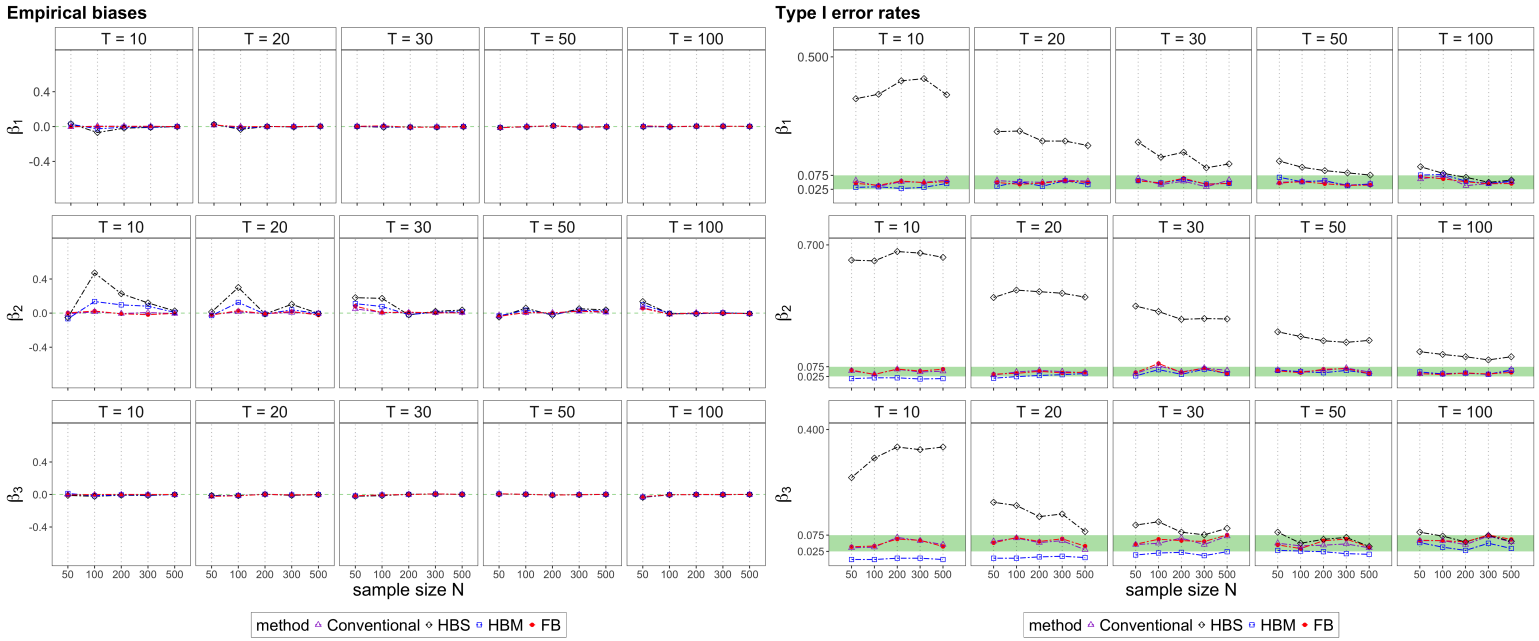
Note. The shaded area marks the boundaries of [.025, .075] for the empirical Type I error rates. Conventional = the conventional approach; HBS = the hybrid-Bayesian approach with a single draw; HBM = the hybrid-Bayesian approach with multiple draws; FB = the one-step full Bayesian approach.

Figure 3. Relative biases and 95% CI coverage rates when all regression coefficients had small-medium effect sizes and IVAR is a predictor



Note. The shaded area marks the boundaries of $[-.10, .10]$ for relative biases and $[.91, .98]$ for 95% CI coverage rates. Conventional = the conventional approach; HBS = the hybrid-Bayesian approach with a single draw; HBM = the hybrid-Bayesian approach with multiple draws; FB = the one-step full Bayesian approach.

Figure 4. Empirical biases and empirical Type I error rates when all regression coefficients were zero and IVAR is a predictor



Note. The shaded area marks the boundaries of [.025, .075] for the empirical Type I error rates. Conventional = the conventional approach; HBS = the hybrid-Bayesian approach with a single draw; HBM = the hybrid-Bayesian approach with multiple draws; FB = the one-step full Bayesian approach.

Table 2

True values, estimates, and 95% CIs of the regression coefficients from the two illustrative examples.

	Predictor	True regression coefficient (β)	Regression coefficient estimate and 95% confidence (or credible) interval			
			Conventional	HBS	HBM	FB
Ex. 1: using ISD	IM	$\beta_1 = 10$	8.26 [6.50, 10.02] *	9.16 [6.81, 11.51] *	9.78 [7.25, 12.30] *	9.93 [7.39, 12.35] *
	AR(1)	$\beta_2 = 8$	6.17 [3.67, 8.68] *	11.22 [7.68, 14.75] *	8.03 [4.14, 11.92] *	7.98 [4.25, 11.51] *
	ISD	$\beta_3 = -2$	1.46 [-.78, 3.71]	-1.95 [-4.44, .53]	-1.61 [-4.19, .97]	-2.00 [-4.51, 0.60]
Ex. 2: using IVAR	IM	$\beta_1 = 10$	9.21 [7.67, 10.76] *	8.63 [6.66, 10.59] *	9.91 [7.64, 12.18] *	9.90 [7.75, 12.04] *
	AR(1)	$\beta_2 = 8$	6.99 [4.41, 9.56] *	12.00 [8.48, 15.53] *	8.74 [5.03, 12.45] *	8.50 [4.56, .12.17] *
	IVAR	$\beta_3 = -2$	-.39 [-1.25, .47]	-1.45 [-2.35, -.55] *	-1.99 [-4.21, .14]	-1.78 [-3.20, -.56] *

Note. * indicates $p < .05$. Conventional = the conventional regression approach; HBS = the hybrid-Bayesian approach with a single draw; HBM = the hybrid-Bayesian approach with multiple draws ; FB = the one-step full Bayesian approach.